DOES MARKET CONCENTRATION MOTIVATE PULP AND PAPER MILLS TO VERTICALLY INTEGRATE?

A Thesis
Presented to
The Academic Faculty

by

Gewei Wang

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the School of Economics

Georgia Institute of Technology

December, 2005

DOES MARKET CONCENTRATION MOTIVATE PULP AND PAPER MILLS TO VERTICALLY INTEGRATE?

Approved by:

Dr. Haizheng Li, Advisor
School of Economics
Georgia Institute of Technology

Dr. Patrick McCarthy
School of Economics
Georgia Institute of Technology

Dr. Vivek Ghosal
School of Economics
Georgia Institute of Technology

Date Approved: August 25, 2005

ACKNOWLEDGEMENTS

I would like especially thank to Haizheng Li for offering me a precious opportunity to enter Tech, continuous guidance on research, and meticulous arrangement for graduation, without his kindness and wisdom I would not go ahead with my ambition in academy. Patrick McCarthy gave me much help on study and graduation. Peter Ince and James McNutt provided insightful discussions for the industry. Jane Chisholm picked hundreds of grammatical errors in this paper. Jifeng Luo and Lei Lei took care of me with endless friendship. Last but not least, I thank to my beloved parents and Meixin.

TABLE OF CONTENTS

LIST OF TABLES

SUMMARY

Following sound economic theory, paper mills vertically integrate into pulp production, partly because internalizing the production of their inputs allows them to avoid transaction costs. Higher market concentration, a proxy of higher asset specificity and transaction costs, should encourage vertical integration in the pulp and paper industry. However, this relationship has not been robust in previous studies or in our replication with updated FPL-UW data. Upon a deeper analysis of the data, this study should clarify the mechanism by which transaction cost can induce vertical integration in this particular industry, which does not have well-defined intermediate goods markets. In order to specify the pulp markets where paper mills are likely to trade, we construct a mill-specific concentration measure as a substitute to traditional regional concentration measures. We also narrow our sample to mills producing free sheet paper, the most profitable paper grade in this industry. With such model refinement, this research exhibits a significantly positive correlation between transaction cost and vertical integration.

JEL Classification: C23, D23, L16, L22, L69

Key Words: Market Concentration, Transaction Cost, Vertical Integration

1. INTRODUCTION

Recent empirical studies on the determinants of vertical integration have witnessed a transition from the use of industry-level data to plant-level data. The accessibility of microdata and the natural dichotomy of the upstream and downstream manufacturers have led to the study of vertical integration in the pulp and paper industry. Ohanian (1994), followed by Melendez (2002), explores the factors that influence the decision by a paper mill to vertically integrate into producing its own pulp and tests the role of transaction cost, which mostly relates to the term of economics among these factors. Transaction Cost Economics (TCE), introduced by Williamson (1975), considers vertical integration as a way of circumventing potential holdup problems. Therefore, internalizing the production of input allows mills to avoid transaction costs. In line with TCE, Ohanian finds that the vertical integration of pulp and paper production is positively associated with regional concentration, an indicator of transaction cost. However, the duplicative model presented in our paper detects results contrary to Ohanian's from the FPL-UW dataset, which records detailed capacities of individual mills from 1970 to 2000.

The positive relationship between transaction costs and vertical integration can neither vanish nor reverse. A possible explanation is that this economic relationship became weaker in contrast to other technological economies, while the pulp and paper industry experienced a rapid evolution after World War II. This paper looks into the new structure of this particular industry and finds that the pulp market that each mill has access to cannot be segregated simply into one of several regions since transportation has improved. If intermediate-good markets cannot be well defined, transaction cost cannot be measured accurately. To construct a concentration measure as a substitute to the traditional regional concentration measures, this paper will define mill-specific markets in which paper mills are more likely to trade.

Differentiated by end use markets as well as by process type, several principal paper grades are recognized in the industry. Among some of the categories of paper production, we find that integrated patterns differ. We select and focus on a sub-sample for paper mills with positive free sheet capacity, thanks to the FPL-UW database. Free sheet paper is virtually true of all printing and writing paper, and thus, is rather expensive compared to other grades of paper, so it may be economical to use market pulp only in free sheet paper. In other words, technological economies dominate transaction-cost economies in some categories of paper production. Therefore, the effect of transaction cost may be hidden behind the results.

This paper is organized as follows. Section 2 introduces the standard transaction cost economics on vertical integration, and Section 3 presents a brief review of the literature that discusses vertical integration in the pulp and paper industry. Section 4 introduces our dataset. Section 5 estimates with a cross-sectional model in previous studies. Section 6 refines our model by investigating the definition of pulp market and a basic sample.

Sections 7 and 8 estimate a model using panel data and check the robustness with various specifications. Section 9 concludes.

2. THE DETERMINANTS OF VERTICAL INTEGRATION

Other than the benefits brought by market imperfection[1], the incentives for vertical integration can be classified into two categories: production-cost reduction and transaction-cost reduction. Production cost, the amount of which varies for each industry or sector, can be retrenched if fewer resources are used to produce the downstream output upon vertical integration with the upstream process. Transaction cost is incurred due to the probability of "opportunistic behavior," which may arise where specific assets are needed. TCE holds that the higher the transaction costs, the fewer the number of parties in the intermediate-goods market and the more specialized the assets involved in the transaction, known as "asset specificity."[2] When assets become specific to a particular transaction, the firm is vulnerable to opportunistic behavior by the supplier (Klein et al., 1978). [3]

---

[1] Such as avoiding supply uncertainty, government intervention, or "double-marginalization".

[2] For example, a desktop computer can be used in many productive activities whereas a Fourdrinier machine can be used only in papermaking.

[3] As the topic of vertical integration has been popular in the domain of new institutional economics for many years, extensive literature reviews on the determinants of vertical integration can be found in recent studies, theoretical reviews such as that in Joskow (2005) and Whinston (2003), and empirical reviews such as that in Klein (2005) and Acemoglu et al. (2004).

3. VERTICAL INTEGRATION IN THE PULP AND PAPER INDUSTRY

Globerman and Schwindt (1986) argue that asset specificity occurs in papermaking because the salvage value for a mill or component assets is low. The fixed-proportions and continuous-process nature of papermaking also imply a greater potential loss from the interruption of input supply compared to variable-proportion or batch-process industries (Ohanian, 1994).

Where asset specificity cannot be easily measured, concentration has been used in some studies to capture situations in which holdup problems are likely to appear; high concentration indicates conditions in which greater potential occurs for exploitation in market transactions and the firm has greater incentive to vertically integrate. MacDonald (1985) shows the levels of vertical integration within the manufacturing sectors appear to be strongly associated with high levels of concentration. Ohanian (1993, 1994) tests a model of vertical integration for the U.S pulp and paper industry between 1900 to 1940, using mill-level data collected from the Lockwood's Directories[4] at ten-year intervals. She finds the vertical integration of pulp and paper production to be positively associated with regional concentration, paper mill capacity, and the production of standardized grades of paper. Melendez (2002) uses the same data source as Ohanian between 1975 and 1995 at five-year intervals. Unlike Ohanian, she finds that vertical integration is

---

[4] *Lockwood's Directory of the Paper and Allied Trades*, published since the 1870's.

positively correlated to paper mill capacity, the production of standardized grades of paper and forestland dummy, but negatively associated with regional concentration. For her effect of market concentration is contrast to that of Ohanian (1994), Melendez argues that market concentration is endogenously determined in the reduced form model, and thus the estimates are biased.

4. DATA

The data (FPL-UW database) are maintained at the USDA Forest Service, Forest Products Laboratory (FPL), in collaboration with the University of Wisconsin-Madison. The data include estimates of annual production capacity for all mill locations in the United States where paper, paperboard, or market pulp was produced between1970 and 2000, our sample period. Records for each mill location and each year are included, and the record includes the company name, the city, the state, the ZIP code, and capacity estimates by process type for each product category[5]. Capacity data at each mill location are further differentiated by process type within each category of paper or paperboard; for example, capacity based on recycled fiber is differentiated from capacity based on woodpulp, and capacity based on integrated chemical pulp is differentiated from capacity based on market chemical pulp.

Other than the common variables, the term "integrated" in the FPL-UW data means that papermaking capacity is combined with pulping capacity at the same facility or mill

[5] The sector is divided into commodity categories within three broad commodity groups: paper, paperboard, and market pulp. The paper commodity group includes eight conventional categories: newsprint, four categories of printing and writing paper (uncoated free sheet, coated free sheet, uncoated groundwood, and coated groundwood), tissue and sanitary paper products, unbleached kraft paper, and other specialty packaging and industrial paper products. The paperboard commodity group includes four conventional commodity categories: linerboard and corrugating medium, solid bleached board, and other recycled paperboard. The market pulp commodity group primarily includes hardwood and softwood kraft market pulp, deinked market pulp based on recycled fiber, and relatively small amounts of bleached chemithermomechanical market pulp (CTMP) and cotton linter pulp. These commodities are produced generally for use in papermaking. For more details, see Ince et al. (2001).

location. Paper and paperboard are downstream goods (thus,“paper and paperboard” will be referred to as “paper” in the following discussion unless otherwise specified), and pulp is upstream or intermediate goods. We want to investigate the intermediate goods markets and the integration of intermediate goods production. Because *Lockwood’s Directory* includes estimates for pulp capacity, Ohanian (1994) and Melendez (2002) view a paper mill as vertically integrated if pulp capacity is a positive value. The FPL-UW data do not have information about pulp capacity, but instead reports the capacity of market pulp that is produced at one location and sold to industrial users at another location or exported. However, although we don’t know the exact pulp capacity for every mill, we can derive the source of pulp from the process of paper production, which means that we can reveal the integration status for each product category.

5. CROSS-SECTIONAL MODEL REPLICATING PREVIOUS STUDIES

Ohanian (1994) studies the transaction-cost economies in vertical integration from the pulp and paper industry. Although this industry evolved throughout the twentieth century, the incentive of vertical integration brought about by transactional economies should not be affected by any changes in industrial structure. From the FPL-UW data, another excellent data source, we want to know what estimates we can get. Therefore, for the first step using our new data, we try to replicate Ohanian's model to see if this model exhibits a positive relationship between market concentration and vertical integration.

Table 1. Statistics of Vertical Integration in the Pulp and Paper Industry, 1975-1995

	1975	**1980**	**1985**	**1990**	**1995**
A. Statistics from the FPL-UW Data					
Number of integrated paper mills	236	233	232	233	222
Number of total paper mills	549	527	505	502	496
Ratio	43%	44%	46%	46%	45%
B. Statistics from the Lockwood's Directory					
Number of integrated paper mills	269	311	246	241	233
Number of total paper mills	584	590	559	508	508
Ratio	46%	53%	44%	47%	46%

Source: the FPL-UW data and Melendez (2002).

Ohanian's logit model is

$$\ln\left(\frac{VI_i}{1-VI_i}\right) = \beta_0 + \beta_1 concentration_i + \beta_2 size_i + X_i'\eta + \varepsilon_i. \tag{1}$$

VI is a dummy variable, equal to one if a paper mill[6] is vertically integrated with the pulping process. The form of integration is often modeled as a discrete variable: "make", "buy", or "hybrid" (Klein, 2005). Ohanian's dependent variable is equal to one if a paper mill has a positive pulp capacity. We define vertical integration directly from the process of paper production: if a paper mill has any one category of paper capacity using market pulp, then it is viewed as an integrated paper mill[7]. Although we use different data and different specifications for vertical integration, the descriptive statistics are similar to those of Melendez (Table 1)[8]. *Concentration*, which refers to regional market concentration, a measure of asset specificity, is defined as the product of the top four mill concentration ratios (CR4) in the pulp (supplier) and paper (producer) markets of each

---

[6] A "paper mill" is defined as "a mill with positive paper and paperboard capacity".

[7] In fact, we don't know the pulp capacity integrated into paper production. What we know is the integrated paper capacity. Generally, the ratio of pulp input and paper output is around one, but we will not boldly derive integrated pulp capacity from integrated paper capacity.

[8] The pulp used in paper production can be classified broadly into three categories: mechanical pulp, chemical pulp, and recycled pulp. Recycled pulp is a bit different from the other two pulping process since a paper mill can purchase recycled (deinked) pulp directly from the pulp market or purchase wasted paper for in-house recycling. Our FPL-UW data does not differentiate the process of recycled pulping clearly. Nevertheless, the industrial profession cannot even tell whether recycled pulping should be called integrated or not. Here, we consider recycled pulping a non-integrated process and find the statistic close to Melendez's estimates, which are derived from *Lockwood's directory*. In the following sections, we try to solve this problem by narrowing the sample.

region[9], based on reported mill capacities[10]. *Size* is actually another proxy of transaction cost because the frequency and volume of transaction rises with mill size[11]. *X* is a vector that includes other mill characteristics. In fact, the specification of Ohanian's model follows:

$$\ln\left(\frac{VI_i}{1-VI_i}\right) = \beta_0 + \beta_1 concentration_i + \beta_2 capacity_i + \beta_3 news_i + \beta_4 kraft_i + \varepsilon_i. \qquad (2)$$

Capacity, the logarithm of the paper capacity of each mill[12], is a proxy of mill size because data such as sales or output are not available in the *Lockwood's Directory* and the FPL-UW database. *News* and *kraft* are two dummies for positive newsprint and kraft capacity, which may represent the higher degree of asset specificity that results from the specialization of pulping assets so that they conform to the requirements of the papermaker. Ohanian's model is estimated with cross-sectional data at ten-year intervals between 1900 and 1940, and correspondingly, we also base our model on the data of the

---

[9] Ohanian's division of four regions is as follows: the North includes CT, DE, ME, MA, MD, NH, RI, VT, NY, NJ, PA; the Lake and Central regions include IL, IN, MI, OH, WI, IA, KS, MN, MO; the South includes FL, GA, NC, SC, VA, WV, AL, MS, TN, AR, LA, TX; the Pacific includes CA, OR, WA. The FPL-UW data include no mills in RI, but they include mill(s) in AK, AZ, CO, DC, ID, KY, MT, NM, OK. We add the states where paper mills were not located before 1940 into the four regions according to the Census division. It should be noted that this division is a little different from that of the FPL-UW data.

[10] Ohanian argues that a regional measure is appropriate because most grades of pulp and paper were traded within the producing region during her sample period.

[11] For details, see Ohanian (1994).

[12] The paper capacity in Ohanian's model is measured by thousands of pounds per 24 hours. We measure the capacity of pulp and paper by thousands of short tons per year. The coefficients of the variables other than *capacity* will not be affected.

sample period at ten-year intervals. Table 2 shows the mean values of the independent variables in our model and of those from Ohanian's Table 1 for comparison. It shows that the pattern of vertical integration changed significantly several decades after 1940. Although the number of integrated paper mills is still lower than the number of non-integrated mills, these numbers became much closer between1970 and 2000. While the total number of paper mills continued to decline during the 20th century, the number of integrated paper mills after 1970 was higher than that before 1940. The ratio of the capacity of integrated paper mills to that of non-integrated paper mills is also higher in our sample period, and the proportion of newsprint and kraft mills declined after the 1940s and stabilized after1970. The last row of Table 2 indicates that the association between market concentration and the status of vertical integration between 1970 and 2000 reversed in contrast to the period between 1900 and 1940. The more concentrated markets are located in regions with less integrated paper mills now.

Table 2. Variable Means for the Model of Paper Mills

Integrated vs. Nonintegrated Paper Mills

	1900		**1940**		**1970**		**2000**	
	NITG	ITG	NITG	ITG	NITG	ITG	NITG	ITG
Number of mills	612	147	546	152	312	239	256	214
Paper capacity†	3	13	22	61	45	173	91	361
Dummy (Newsprint)	0.06	0.37	0.00	0.16	0.01	0.08	0.02	0.09
Dummy (Kraft)	—	—	0.07	0.26	0.02	0.12	0.03	0.07
Concentration (CR4)	0.04	0.06	0.05	0.07	0.09	0.08	0.12	0.09

Notes: †. Thousand short tons per year.

1. NITG: non-integrated; ITG: integrated.

Source: the FPL-UW data and Ohanian (1994).

Table 3. Logit Regression for Paper Mills, 1970-2000

Basic model with the sample of total paper mills

	1970	**1980**	**1990**	**2000**
Concentration (CR4 prod.)	-14.968***	-7.239**	-5.010**	-2.123
	(4.006)	(3.145)	(2.479)	(1.756)
Capacity (log)	0.748***	0.799***	0.787***	0.779***
	(0.094)	(0.103)	(0.101)	(0.098)
Dummy (Newsprint)	1.325**	1.209*	0.712	1.085*
	(0.662)	(0.653)	(0.587)	(0.583)
Dummy (Kraft)	1.177**	1.323**	0.622	0.645
	(0.531)	(0.544)	(0.543)	(0.547)
Constant	-2.029***	-3.020***	-3.196***	-3.684***
	(0.549)	(0.609)	(0.592)	(0.557)
Log likelihood	-299.82	-285.64	-281.64	-266.07
N	551	527	502	470

Note: 1. The dependent variable, vertical integration, is a dummy variable, indicating a paper mill is vertically integrated if it is integrated with the pulping process. The concentration ratio is equal to the product of the regional top-four paper mill capacity ratio and regional top four market pulp capacity ratio.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

The estimates of Model (2) are reported in Table 3. All paper mills with positive paper and paperboard capacity are included in each regression. Because the number of mills operating each year varied with entry and exit, the number of observations ranges from 470 to 551 over the period.

The integration status is shown to be positively associated with mill capacity and the two dummies of specialized papermaking over our sample period as expected. The coefficients of mill size are statistically significant in each sample year, but the coefficients of specific paper grades are insignificant in some years after the 1980s. The surprising result from this exercise is that the coefficients of concentration measure are negative and significant, which clearly contradicts Ohanian's estimates. Melendez (2002) also fails to obtain same explanation of transaction cost as Ohanian with newly updated data from the *Lockwood's Directory* between 1975 and 1995. She explains that the cause of this discrepancy is the presence of endogeneity in the model. During the first half of the century, the paper industry in the U.S. was going through a relocation process from the North towards the South by acquiring large timber tracts and constructing high-capacity pulp mills[13]. Meanwhile, the company managers deemed forward integration into papermaking necessary for sustained competitiveness (Toivanen, 2004). The number of

[13] Toivanen (2004) details the reason as, "Rapid diffusion of kraft pulp innovations propelled industrial relocation in the North American pulp and paper industry. The cost and availability of pulp wood in traditional industrial regions was a persistent subject of debate. ...... The kraft process was the first pulping technology that could potentially utilize on large scale the vast stands of Southern pine, often characterized as 'waste,' and thus prompted managers to consider Southern strategy."

mills in the South was small, indicating a high concentration, but the ratio of integrated paper mills was higher in the South than that in the North. This relationship is revealed in Ohanian's model, which is consistent with the transaction cost theory. After the 1970s, the industry settled, and the differences in market concentration across regions narrowed. However, the ratio of the proportion of vertically integrated paper mills in the South and to those in the North remained the same due to the natural advantage of resources in the South.

The fact that Ohanian's model fails to detect the relationship between transaction cost and vertical integration suggests that some unobservable factors exist. One of the most important but difficult factors to control is forest resources. If timber tracts are abundant in one area, pulp mills are likely to be established, which results in a low seller- market concentration. For the same reason, paper mills located there are also likely to be integrated with the pulping process. As a result, a high level of vertical integration correlates with a low level of market concentration unless all important factors are controlled. Since technical economies dominate other factors affecting integration decisions in this particular industry, the effects of other contributing factors, such as transaction cost, are less likely to be correctly detected. We know if an omitted variable is positively correlated with the dependent variable and negatively correlated with the independent variable, downward bias of the estimates occurs. In fact, our estimates for concentration from the FPL-UW data are significantly negative, which seems to be strongly biased. Melendez proposes a dynamic model that explores the likelihood of endogeneity and corrects these biases with some simulation exercises. Diverting from the application of structural models, we delve into the problem of endogeneity by refining specifications and estimation.

6. REFINEMENT IN SPECIFICANTIONS

Lagged explanatory variables

The choice of either integrating pulping capacity or purchasing market pulp is usually a long-term decision for a mill. Therefore, we replace the concentration measure with a lagged value that also reflects the possibility that a company manager is not aware of the market situation in the current year. The concentration measure that we choose is a three-year-lagged CR4 product. Due to a high correlation among concentration measures across years, the results differ little from those of the model with current-year concentration measures. Using lagged variables per se cannot solve the endogeneity problem. We will keep using lagged variables for market concentration in that they make sense for formulating strategy on vertical integration.

A sample based on free sheet mills

Technology economies in the pulp and paper industry lies primarily in reducing processing costs involved in producing market pulp versus pulp used directly at an integrated pulp and paper mill[14]. On the other hand, the cost advantage to integrated paper mills varies, depending on the structure and sales price of the final product (Zavata, 1993). The savings from integration are approximately five to ten percent of pulp manufacturing costs (Diesen, 1998) and can be as high as fifteen percent, but they are much less

---

[14] Typically, market pulp must be dried and baled before being shipped to the paper mill, where it has to be diluted again. In an integrated paper mill, wet pulp is sent directly to the paper machine. The elimination of the pulp-drying stage reduces both investment and operating costs, because a mill no longer needs for a drying machine, and saves energy and packaging material costs (Zavata, 1993).

significant in the case of high value-added paper. Free sheet paper, which accounts for all fine writing and printing paper, is produced primarily with bleached chemical pulp, with generally less than ten percent mechanical pulp content[15]. Even though the cost of chemical pulp is much higher than that of recycled pulp and mechanical pulp, free sheet paper is the most profitable because it is relatively expensive, compared to other grades of paper with low-content of chemical pulp[16]. As a result, almost all market pulp produced and sold in North America is bleached kraft pulp. Hardly any mechanical pulp nor unbleached kraft pulp is sold in the market[17]. In other words, mechanical pulping is naturally integrated in most mills. Therefore, papermaking with mechanical pulp is integrated where a mill is built. Today many newsprint, kraft and paperboard mills are able to eliminate the use of purchased market pulp because of improvements in technology.

---

[15] For the pulp categories, the following common divisions are used: mechanical pulp (literally, pulp produced by a mechanical process), chemical pulp (literally, pulp produced by a chemical process), semi-chemical pulp (as the name suggests, pulp made by a chemical process followed by a mechanical process), and other pulp (for example, deinked pulp or recycled pulp).

[16] For example, coated free sheet (No. 1 grade) currently sells for about $1,500 per ton, whereas newsprint (an uncoated paper product made from high-yield mechanical pulp, often with a very small fraction of bleached chemical pulp) currently sells for about $500. Bleached kraft market pulp costs about $300 to $400 per ton to make, so it may be economical to use market pulp in coated free sheet paper, but not in newsprint (the price of newsprint is too low to support the use of market pulp).

[17] In 2000, the production of mechanical pulp in the U.S. was 1,745 tonnes compared with that of chemical pulp at 48,199 tonnes. Import of mechanical pulp in the U.S. was as low as 437 tonnes. In our capacity estimates from the FPL-UW data, the capacity of U.S. market pulp was zero in the nine years between 1970 and 2000.

In the FPL-UW database, mechanical pulping and unbleached chemical pulping processes are not shown whether purchased or integrated. Only three paper categories (uncoated free sheet, coated free sheet, and tissue), which mostly use chemical pulp, are likely to purchase market chemical pulp. In tissue papermaking, recycled pulp, used extensively, cannot easily be identified as integrated or non-integrated. After all, recycled pulping may be carried out by either de-inking recovered paper or purchasing market recycled pulp directly. Although recycled pulping is generally referred to as a non-integrated pulping process, the real industrial scenario is complicated, so we do not want to address this issue. Therefore, this paper focuses on two categories of paper production: coated free sheet paper and uncoated free sheet paper[18]. Free sheet mills (referred to as mills with positive free sheet capacity) only need bleached chemical pulp, so they are more likely to involve a decision of whether to vertically integrate or non-integrate.

Table 4 presents some summary statistics. The size of free sheet mills grew and the number decreases between 1970 and 2000, which reflects the tendency of the entire industry. The total capacity of free sheet mills increased 145 percent between 1970 and 2000, while the number of free sheet mills decreased from 82 to 80, and the number of free sheet mills decreased from 116 to 108, mostly after the 1990s. The number of firms and single-mill firms (referred to as firms with one mill only) declined even faster at

[18] Coated free sheet paper generally falls on the highest value end of the printing and writing paper spectrum and used almost entirely in commercial printing applications such as annual reports, product sales brochures, or advertising pamphlets that generally demand high image quality and color printing. Uncoated free sheet paper is primarily used for producing office reprographic paper for copies and printers.

nearly 30 percent, caused by the trends of horizontal mergers and mill exits, which led to a more concentrated buyer market for pulp. As for vertical integration, Table 5 shows that integration ratios rose from 56 to 64 percent, a moderate increase, during these four sample years. The ratio between the paper capacity of integrated and non-integrated mills gradually increased from 60 to 75 percent during the thirty years.

Table 4. Descriptive Statistics for Free Sheet Mills, 1970-2000

	1970	**1980**	**1990**	**2000**
Number of free sheet mills†				
Free sheet capacity only	82	87	88	80
Total	116	117	120	108
Number of free sheet firms	50	48	36	35
Single free sheet mill firms†	30	30	22	22
	60%	63%	61%	63%
Total free sheet capacity (thou. short tons/year)	9022	12128	16808	22069

Notes: †. Free sheet mills are defined as mills with positive free sheet capacity. Single free sheet mill firms denote firms with only one free sheet mill.
Source: the FPL-UW data.

Table 5. Integration Statistics for Free Sheet Mills, 1970-2000

	1970	**1980**	**1990**	**2000**
Integrated mills	65	69	74	69
Total mills	116	117	120	108
Ratio	56%	59%	62%	64%
Integrated capacity	5408	7822	11703	16438
Total capacity	9022	12128	16808	22069

Source: the FPL-UW data.

Reconsiderations in the pulp market

Most grades of pulp and paper were traded within the producing regions at the beginning of the 20th century, which motivates Ohanian (1994) to use a regional concentration measure. Nowadays, market pulp can be transported by sea or by rail a very long distance at much lower costs than before[19]. The import of chemical pulp in the U.S. is equal to thirteen percent of chemical pulp production in 2000. Hence, assuming that pulp is traded only within regions seems not in tune with the times. On the other hand, another kind of pulp called “wet-lap pulp”, which is not totally dried and has lots of moisture (about fifty percent), is still traded in the pulp market, albeit in small quantities. While wet-lap pulp is more cost-effective, it tends to deteriorate rapidly if it remains wet for a very long time. However, since highway transportation, a common mode of transportation, is efficient in customer service, it can ensure higher quality of the pulp bundles, integrity of its

[19] For ship and rail transport, loading and unloading costs are significant, but transit costs are relatively lower, so once the goods are loaded onto a ship or into a railcar, shipping distances are not so critical.

packaging and reduction of costs. A paper mill can utilize trailers with a payload capacity of 30 tons, dedicated 24 hours per day to pulp transportation. Therefore, according to industry sources and recommendations of Melendez (2002), we define a circular market for each paper mill of a 350-mile radius, in which the paper mill can purchase market pulp at lower costs. We refer to this market as a "neighbor market" in contrast to a "peripheral market," which includes the U.S. as well as foreign countries outside the circular area. Such an explicit market boundary has been specified a bit arbitrarily; however, according to robustness checking, adjustment within 100 miles of the radius can hardly alter the estimated results. This distance is considered the maximum transportation distance for wet-lap pulp and an effective distance of highway transportation for dried market pulp. Due to the likelihood that no mill will be located within the circular market of some paper mills, we have added the five closest pulp or paper mills outside the "neighbor market" in the 350-mile radius circular market for each mill to form seller markets and buyer markets. This setup is useful, as it reflects the relative accessibility and competition for a paper mill to purchase market pulp.

In contrast to the CR4 index, our mill-specific concentration measure, which allows us to further control the regional characteristics and some endogeneity by adding dummies or using panel regression, is based on the Herfindahl-Hirschman Index (HHI), one commonly used measure of market concentration. We separately calculate the HHI for the buyer market (HHI paper) and the seller market (HHI pulp) with the market definition introduced above. It should be noted that within the neighbor market of each paper mill, only paper mills or pulp mills[20] not belonging to the same firm are included to calculate this mill-specific HHI, so it is a measure of firm concentration. By multiplying these two concentration measures, we get a variable of HHI interaction, which we will also use in our models.

Concentration of the seller market may exert distinct effects on integration strategy for some pulp mills under the same ownership in the "neighbor market" (referred to as "sister pulp mills") and for the mills that don't have sister pulp mills and are thus unlikely to take advantage of in-firm pulp purchasing. We expect that the concentration will not correlate with vertical integration if a mill has access to the pulp from its own firm. Vertical integration cannot be defined on the firm level in the pulp and paper industry, as distance and mill distribution should be taken into account. Therefore, our definition of mill integration is not the same as that of firm integration, and we add a dummy variable of a sister pulp mill in a neighbor market to reflect the relationship between these two levels of vertical integration. A dummy variable indicating whether a mill is located in the South is also added in order to represent the regional difference in timber tracts now that we have a more specific concentration measure.

[20] Pulp mill is referred to as a mill with positive market chemical pulp capacity.

Estimates

Table 6. Variable Means for the Model of Free Sheet Mills

Integrated vs. Nonintegrated Free Sheet Mills

	1980		1990		2000	
	NITG	ITG	NITG	ITG	NITG	ITG
Number of mills	48	69	46	74	39	69
Paper capacity†	58	136	77	179	75	278
Dummy (uncoated free sheet capacity)	0.833	0.942	0.826	0.892	0.872	0.855
Dummy (south)	0.021	0.246	0.043	0.311	0.051	0.319
Dummy (sister pulp mills)	0.146	0.217	0.196	0.230	0.051	0.261
HHI (product) 3-year-lagged	1190	2279	1963	2288	2558	2907
HHI (seller market) 3-year-lagged	1433	1608	1784	1711	2011	2024

Notes: †. Thousand short tons per year.

1. NITG: non-integrated; ITG: integrated.

Source: The FPL-UW data.

Table 7. Logit Regression for Free Sheet Mills, 1980-2000

Cross-sectional model with HHI product as concentration measure

	1980	1990	2000
HHI (product) 3-year-lagged	0.338	0.028	-0.108
	(0.273)	(0.124)	(0.171)
Capacity (log)	1.236***	1.051***	1.560***
	(0.280)	(0.263)	(0.331)
Dummy (uncoated free sheet capacity)	1.498**	0.886	0.693
	(0.754)	(0.628)	(0.758)
Dummy (south)	1.874	1.659**	1.146
	(1.223)	(0.812)	(0.935)
Constant	-6.745***	-5.241***	-7.094***
	(1.446)	(1.403)	(1.842)
Log likelihood	-53.70	-62.07	-46.09
N	117	120	108

Notes: 1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI product is a concentration measure obtained by multiplying the HHI of the seller and buyer markets for chemical pulp.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

Table 8. Logit Regression for Free Sheet Mills, 1980-2000

Cross-sectional model with HHI of seller market as concentration measure

	1980	**1990**	**2000**
HHI (seller market) 3-year-lagged	0.522	-0.349	0.672
	(0.735)	(0.616)	(0.579)
Capacity(log)	1.281***	1.065***	1.560***
	(0.276)	(0.262)	(0.338)
Dummy (uncoated free sheet capacity)	1.642**	0.950	0.674
	(0.770)	(0.631)	(0.766)
Dummy (south)	2.210*	1.560*	1.384
	(1.168)	(0.832)	(0.975)
Constant	-7.340***	-4.670***	-8.791***
	(1.830)	(1.681)	(2.327)
Log Likelihood	-54.93	-61.93	-45.60
N	117	120	108

Notes: 1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI of the seller market is a concentration measure for the chemical pulp market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

Table 9. Logit Regression for Free Sheet Mills, 1980-2000

Cross-sectional model with estimation of two separate coefficients for the HHI of the seller market as the concentration measure

	1980	**1990**	**2000**
HHI-WS (seller market) 3-year-lagged	1.558	-3.054**	1.132
	(1.510)	(1.371)	(4.605)
HHI-NS (seller market) 3-year-lagged	0.188	1.303	0.666
	(0.771)	(1.055)	(0.596)
Dummy (sister pulp mills)	-2.177	7.470**	-0.780
	(2.549)	(3.107)	(7.674)
Capacity (log)	1.349***	1.075***	1.563***
	(0.295)	(0.276)	(0.357)
Dummy (uncoated free sheet capacity)	1.718**	0.939	0.676
	(0.781)	(0.676)	(0.787)
Dummy (south)	2.179*	1.984**	1.405
	(1.178)	(0.895)	(1.005)
Constant	-7.144***	-7.559***	-8.794***
	(1.777)	(2.255)	(2.349)
Log Likelihood	-54.56	-57.55	-45.60
N	117	120	108

Notes: 1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has a sister mill within its 350-mile radius circular neighbor market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

Table 7, 8, and 9 show some cross-sectional estimates for free sheet mills and variable means are shown in Table 6. Three common independent variables are added in these models plus a different concentration measure. Free sheet capacity reflects mill size, a proxy of asset specificity, the dummy of uncoated free sheet capacity indicates the product difference in the requirement of vertical integration, and the dummy of the South region is an indicator of forestland abundance. The estimates of these three variables are robust in different models shown in Table 7, 8 and 9. The coefficients of free sheet capacity are always significantly positive. The positive estimates of uncoated free sheet dummy show evidence that the lower the paper grade is, the more likely it is to be vertically integrated. The dummy of the South always positively correlates with the status of vertical integration, but it is not significant in some years, which can not be detected from the statistics in Table 6. The product of the HHI in the buyer and seller markets is applied in the model of Table 7. The concentration measure exhibits positive effects in 1980 and 1990, albeit insignificant from zero. MacDonald (1985) uses the HHI for the buyer and seller markets separately and finds that the concentration of the buyer and seller markets both positively correlate with the level of vertical integration. He also tests for the interaction of these two concentration measures but finds no strong evidence of interaction effects between buyer and seller concentration.

The buyer markets of trading pulp are much less concentrated than the seller markets. Since more than five free sheet mills are usually located with the 350-mile circular markets[21], they can be considered more competitive. Therefore, we prefer to use concentration of the seller market only, which we consider it more related to transaction

[21] In 2000, 97 percent of free sheet mills have more than five other free sheet mills located in their circular neighbor markets and 85 percent of them have more than ten competitors in the circular markets. On the contrary, the corresponding proportions for free sheet mills with more than five and ten chemical pulp mills in their circular markets are 26 percent and 15 percent respectively. Although we don't consider ownership here, the difference between seller market and buyer market for chemical pulp is significant.

costs incurred in trading pulp. In Table 8, concentration is measured by the HHI of the seller market. The coefficients of concentration are positive, but still insignificant.

As we discussed above, free sheet mills with or without sister pulp mills in the neighbor market may have different strategies of integration. In Table 9, two coefficients for concentration are estimated for mills with sister pulp mills and mills without sister pulp mills in the neighbor market, referred to as HHI-WS and HHI-NS[22] respectively. In addition, a dummy variable indicating whether any sister pulp mill exists in the neighbor market is included. However, the results are still not acceptable. In the cross-sectional years of 1980 and 2000, two coefficients of concentration are both positive but insignificant, and the coefficients of the sister pulp mill dummy is negative. In the model for 1990, the coefficient of HHI-WS becomes negative, and the estimate of the sister mill dummy is positive.

We still have an option of using all the sample years to run a pooled regression. As the sample size increases, more significant results can be obtained. In the first three columns in Table 10, we use similar specifications to those in the models in Table 7, 8 and 9 except we add the dummy of the sister pulp mill into the model of every column. In column I, whereas the concentration shows an effect significantly positive at 10% level,

[22] HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for the isolated mill without sister mills in the neighbor market. These two variables are constructed by multiplying HHI with relevant dummies.

Table 10. Logit Regression for Free Sheet Mills, 1973-2000

Pooled cross-sectional model with various specifications of concentration measure

	I	II	III	IV
HHI (product)†	0.034*			
	(0.023)			
HHI (seller market) †		-0.059		
		(0.095)		
HHI-WS (seller) †			-1.016***	-0.983***
			(0.237)	(0.238)
HHI-NS (seller) †			0.156	0.191*
			(0.109)	(0.112)
Import intensity†				0.017*
				(0.011)
Capacity (log)	1.265***	1.273***	1.248***	1.255***
	(0.053)	(0.053)	(0.053)	(0.054)
Dummy (sister pulp mills)	-0.309***	-0.285**	1.657***	1.662***
	(0.116)	(0.115)	(0.448)	(0.448)
Dummy (UCFS capacity)	1.089***	1.096***	1.071***	1.067***
	(0.129)	(0.129)	(0.130)	(0.130)
Dummy (south)	1.700***	1.717***	1.740***	1.746***
	(0.176)	(0.176)	(0.175)	(0.176)
Constant	-6.166***	-6.044***	-6.298***	-6.958***
	(0.279)	(0.316)	(0.324)	(0.546)
Log Likelihood	-1572.02	-1572.96	-1562.37	-1561.21
N	3,235	3,235	3,235	3,235

Notes: †. Three-year lagged variable.

1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI product is a concentration measure by multiplying the HHI of the seller and buyer markets for chemical pulp. The HHI of the seller market is used as another concentration measure for the chemical pulp market. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has a sister mill within its 350-mile radius circular neighbor market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

which we have never seen so far, the coefficient sign of the dummy of the sister pulp mill is wrong. Column III seems to indicate a result that is closer to our expectation, as the coefficient of the sister pulp mill is significantly positive, and the two coefficients of concentration show different signs. It is confusing that the coefficient of HHI-WS is significantly negative and the coefficient of HHI-NS is insignificantly positive, which may still indicate the existence of downward biases. A new variable, import intensity, is added in the model of column IV. Although we construct a proxy of transaction cost in the neighbor market, the transaction cost in the peripheral market and world market cannot be identified easily. Import intensity is calculated as the market pulp import divided by total sales for each year in the United States. We expect that transaction costs will be higher if the import intensity is greater because high demand for imports leads to opportunistic seller behavior. The estimates of other variables being robust, the coefficient of the HHI-NS becomes significantly positive at a ten percent level compared with that in column III. Import intensity exhibits a positive relationship with vertical integration, also significant at a ten percent level.

7. PANEL MODELS

So far, we have found no strong evidence that market concentration and vertical integration are positively related as indicated by standard theory. Due to unobservable factors, we can control time-invariant factors that are unobservable or unidentifiable, such as local resources, by extending our analysis to panel regression, which has become feasible due to recent development of panel model for binary choice[23] (e.g., see the summary in Green, 2003). Following the analysis in last section, this section will focus on free sheet mills.

Fixed effect model

When we assume the individual-specific effect is non-stochastic, the fixed effect model can be applied in the following form:

$$\ln\left(\frac{VI_{it}}{1-VI_{it}}\right) = \alpha_i + \beta_1 concentration_{it} + \beta_2 capacity_{it} + X'_{it}\eta + \varepsilon_{it}. \tag{3}$$

VI, *concentration*, and *capacity* are defined in the same way as model (2). α_i denotes an individual-specific effect, and *X* includes the dummies of the sister pulp mill and uncoated free sheet capacity. Columns I to IV of Table 11 report the estimates of the fixed effect models. In column I, a three-year lagged HHI of the seller market is applied. Compared to the cross-sectional model of column I in Table 10, the positive coefficient

[23] Our panel data are unbalanced because not all mills were in operation in all thirty years. Some mills entered the industry after 1970, and some exited. Unbalanced panel data still work for the fixed effect and random effect models.

Table 11. Panel Logit Regression for Free Sheet Mills, 1973-2000

Fixed effect models

	I	II	III	IV	V
HHI (seller) †	1.896***				
	(0.615)				
HHI-WS (seller) †		1.889***	-1.436	-0.272	0.002
		(0.679)	(1.459)	(1.858)	(0.019)
HHI-NS (seller) †		1.824***	2.662***	4.946***	0.052***
		(0.603)	(0.730)	(1.239)	(0.009)
Capacity (log)	5.136***	5.186***	4.999***	13.297***	0.056***
	(0.919)	(0.919)	(0.928)	(2.679)	(0.009)
Dummy (sister pulp mills)	0.622		10.137***	10.436**	0.091***
	(0.688)		(3.893)	(4.987)	(0.034)
Dummy (UCFS capacity)	4.267***	4.250***	4.580***	3.880	0.148***
	(1.183)	(1.186)	(1.201)	(2.547)	(0.021)
N (Groups)	75	75	75	75	75

Notes: †. Three-year lagged variable.

1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI of the seller market is used as a concentration measure for the chemical pulp market. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has a sister mill within its 350-mile radius circular neighbor market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

3. Column I : fixed effect model
Column II : fixed effect model
Column III : fixed effect model
Column IV : two way fixed effect model with time effects
Column IV : fixed effect model (linear probability model)

Source: the FPL-UW data.

of concentration is more significant, now at a 1% level. Furthermore, the coefficient of the dummy of the sister pulp mill becomes positive, but not statistically significant. Downward biases were adjusted to some extent, but still exist observably. We try to obtain two separate estimates for the concentration of the seller market in the model of column II and drop the sister mill dummy temporarily, and find that seller market concentration significantly affects both types of mills positively, indicating that if integration is not controlled, concentration has a delusive relationship with integration due to the positive correlation between concentration in the neighbor seller market and the local existence of sister pulp mills. When the sister mill dummy is added in the model of column III, the result is perfectly consistent with what we expected: the coefficient of HHI-WS is insignificant from zero, and the coefficient of HHI-NS is significantly positive. The dummy of the sister pulp mill also shows a significantly positive effect, with z-statistics more than 2.6. In other words, if a mill has some sister pulp mills nearby, it is more likely to be integrated than isolated mills, given that the seller concentration in a neighbor market maintains constant, and no correlation exists between seller concentration and integration strategies for mills with access to in-firm purchasing. The model of column IV is a two-way panel model with estimation of time effects for each year. The addition of time effects does not change the estimated effects of HHI-WS and HHI-NS. We find a significantly positive time trend in most years before 1990. However, in the last decade of these thirty years, the annual time effects on vertical integration for free sheet mills are insignificantly positive. Some studies on vertical integration also use linear probability models with panel data (e.g., Acemoglu et al., 2004). The following equation is such a linear probability model with fixed effects:

$$VI_{it} == \alpha_i + \beta_1 concentration_{it} + \beta_2 capacity_{it} + X_{it}'\eta + \varepsilon_{it}. \qquad (4)$$

The results are shown in column V of Table 11. The estimates of column V become more significant than those in the column III with the exception of the effect of HHI-WS, which is highly insignificant from zero. Therefore, the relationship between concentration and vertical integration is further verified.

Other models of panel regression

After controlling the idiosyncratic time-invariant factors, we find a significantly positive relationship between market concentration and vertical integration. This relationship holds for mills with sister pulp mills nearby because mill integration is essentially determined by firm integration. In order to check the robustness of the fixed effect model, the conditional fixed effect model and the random effect mode are tested.

An incidental parameters problem occurs in the fixed effect model for a binary dependent variable (e.g., Greene, 2003 and Hsiao, 2003), which means that as the time length (*T*) of panel data is typically limited, even when the number of individual (*N*) tends to infinity, the maximum likelihood estimators of all the parameters remains inconsistent. The maximum length of time period in our sample is 28[24], considered moderately large.

[24] As three-year lagged concentration is used, our sample years are from 1973 to 2000.

Table 12. Panel Logit Regression for Free Sheet Mills, 1973-2000

Conditional Fixed effect models

	I	II
HHI (seller) †	2.081***	
	(0.577)	
HHI-WS (seller) †		-1.344
		(1.404)
HHI-NS (seller) †		2.808***
		(0.691)
Capacity (log)	4.858***	4.722***
	(0.869)	(0.874)
Dummy (sister pulp mills)	0.648	10.311***
	(0.679)	(3.749)
Dummy (UCFS capacity)	4.225***	4.528***
	(1.157)	(1.171)
N (Groups)	139	139

Notes: †. Three-year lagged variable.

1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI of the seller market is used as a concentration measure for the chemical pulp market. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has sister mills within its 350-mile circular neighbor market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

Chamberlain (1980) presents a conditional fixed effect model in which estimates are less efficient than they are in the fixed effect model, but consistent even if *T* does not tend to infinity. Table 12 shows estimates from two conditional fixed effect models. The results are quite similar to those in Table 11. While one common coefficient of concentration is estimated (the column I), this coefficient is significantly positive. However, the coefficient of the dummy for the sister pulp mill is insignificant. When two coefficients of concentration are estimated, the one of HHI-WS is insignificant, and the other one of HHI-NS is significantly positive. Meanwhile, the coefficient of the dummy for the sister pulp mill becomes highly significant.

Associated with the fixed effect model, the random effect model can be used under the assumption that the idiosyncratic time-invariant factors are independent of the capacity and market concentration. Table 13 presents the estimates of the random effect models, which slightly differ from those in the fixed effect models. In the one-concentration-coefficient model (column I), the coefficient of concentration is insignificantly positive, and the coefficient of the dummy for sister pulp mill is insignificantly negative. However, with separate concentration effects applied (column II), the results are again consistent with what we expected. The coefficient of HHI-WS is insignificant from zero, and the coefficient of HHI-NS is significantly positive. The coefficient of the dummy for the sister mills is still positive, but only significant at a 30% level, which needs to be further investigated. We can add time-invariant variables in the random effect model. When the

Table 13. Panel Logit Regression for Free Sheet Mills, 1973-2000

Random effect models

	I	II	III
HHI (seller) †	0.037 (0.082)		
HHI-WS (seller) †		-0.749 (1.016)	1.975** (0.945)
HHI-NS (seller) †		0.274*** (0.086)	3.124*** (0.186)
Capacity (log)	2.304*** (0.105)	3.101*** (0.133)	3.258*** (0.200)
Dummy (sister pulp mills)	-0.095 (0.197)	2.271 (2.235)	3.752* (2.108)
Dummy (UCFS capacity)	0.431 (0.357)	0.966** (0.484)	4.874*** (0.915)
Dummy (south)			4.281*** (1.462)
Constant	-7.698*** (0.668)	-12.527*** (0.855)	-21.836*** (1.656)
Random Effect	7.070*** (0.293)	6.917*** (0.289)	7.341*** (0.518)
Rho	0.938*** (0.006)	0.936*** (0.006)	0.942*** (0.009)
N (Groups)	139	139	139

Notes: †. Three-year lagged variable.

1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI of the seller market is used as a concentration measure for the chemical pulp market. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has sister mills within its 350-mile radius circular neighbor market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

South dummy is added (column III), all coefficients are significantly positive, at least at a 10% level. The random effect models do not show as robust results as the fixed effect models, but they are still acceptable and close to our expectation.

The variables we introduce into the models may not capture all the determinants of vertical integration, but we find it difficult to locate other panel variables in the FPL-UW or outer sources. We add an import intensity variable to our panel models to capture the global effect of transaction cost. Import intensity, a time-series variable, has only one value for each year. Adding such a factor cannot bring a better fit for the two-way fixed effect model with full year dummies, but this is a good attempt to the random effect model. The estimated results are shown in Table 14 with the random effect model in column I and the fixed effect model in column II. The coefficients of import intensity are significantly positive and consistent with our expectations. If the import intensity is higher, the domestic demand must increase and transaction cost will be higher if the foreign supply does not vary to some extent. The coefficient of HHI-WS becomes insignificant compared to that estimated without time factors in column III of Table 11, the same as that in the fixed effect model now.

Table 14. Panel Logit Regression for Free Sheet Mills, 1973-2000

Import intensity added

	Random effect model	Fixed effect model
HHI-WS (seller) †	-0.380	-1.163
	(1.019)	(1.522)
HHI-NS (seller) †	1.991***	2.368***
	(0.222)	(0.741)
Capacity (log)	5.537***	6.027***
	(0.267)	(1.099)
Dummy (sister pulp mills)	7.025***	8.781**
	(2.151)	(4.012)
Dummy (UCFS capacity)	3.195***	4.356***
	(0.661)	(1.271)
Dummy (south)	11.127***	
	(1.227)	
Import intensity†	0.070***	0.134***
	(0.019)	(0.048)
Constant	-34.084***	
	(2.001)	
Random Effect	6.392***	
	(0.324)	
Rho	0.925***	
	(0.007)	
N (Groups)	139	139

Notes: †. Three-year lagged variable.

1. The dependent variable, vertical integration, is a dummy variable, indicating a free sheet mill is vertically integrated if it is integrated with the chemical pulping process. The HHI of the seller market is used as a concentration measure for the chemical pulp market. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has sister mills within its 350-mile radius circular neighbor market. Import Intensity is calculated as the market pulp import divided by total sales for each year in the United States.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

8. ROBUSTNESS CHECK FOR INTEGRATION DYNAMICS

In the above sections, we have focused on analyzing the relationship between integration status and market concentration. Integration dynamics is more interesting because it implies more causal interpretation for independent variables. Since the number of mills in this industry keeps declining, the size of the sample of new-entry free sheet mills is relatively small. We are considering the sample in which free sheet mills increase their free sheet capacities compared with the previous year, which also includes all the new-entry mills. The sample size is thus enlarged from 29 to 648, but the sample is notably unbalanced. Although some mills continued to increase their capacities in recent years, capacity increase occurred at some mills only one time, so these mills were dropped in the fixed effect models. Therefore, we can give the estimates only from pooled models. Some variables are redefined in the pooled models for the new sample. The binary dependent variable indicates whether the increased capacity is vertically integrated or not compared to previous year[25]. The variable of capacity measures increased capacity. The dummy variable of uncoated free sheet capacity indicates whether increased capacity included uncoated free sheet capacity or not. In addition, the concentration variable is one-year lagged since capacity increase can be a short term strategy in contrast to the

[25] In the case of capacity increase, we know the change in integrated and non-integrated capacities. It is likely that total capacity increases while one type of capacity (integrated or non-integrated) decreases. Please see Table 15 for an illustrative explanation. We denote an integrated capacity increase if the integrated capacity increases, and the capacity increase is called a non-integrated capacity increase otherwise. The case in which both integrated and non-integrated capacity increases is viewed as an integrated capacity increase.

change in total capacity. Obviously, the estimates for integration in capacity increase (Table 16) are very similar to the estimates obtained from the integration status model (Table 10) that examines vertical integration in total capacity. If a single effect of concentration is estimated, this effect is insignificantly positive regardless of whether or not the sister pulp mill dummy is added (columns I and II). If two separate effects of concentration are estimated, the estimates are similar to those estimated in the pooled logit model for integration status (column III and IV). The coefficient of HHI-WS is significantly negative while the coefficient of HHI-NS is significantly positive. Since we cannot go further to the panel models, the robustness in pooled models cannot guarantee the same results in the panel models. However, as the estimates are biased downward in the pooled models, we anticipate a more significant effect of HHI-NS in the correctly specified model, which is consistent with the TCE.

Table 15. Definition of Vertically Integrated Capacity Increase

	1	**2**	**3**	**4**	**5**
	Integrated Capacity Increase			Non-integrated Increase	
Integrated capacity	↑	↑	↑		↓
Non-integrated capacity	↓		↑	↑	↑

Notes: In the case of capacity increase, we know the change in integrated and non-integrated capacities. It is likely that total capacity increases while one type of capacity (integrated or non-integrated) decreases. We denote an integrated capacity increase if the integrated capacity increases, and the capacity increase is called a non-integrated capacity increase otherwise. The case in which both integrated and non-integrated capacity increases is viewed as an integrated capacity increase (column 3).

Table 16. Logit Regression for Free Sheet Mills, 1973-2000

Pooled cross-sectional model for vertical integration in capacity increase

	I	II	III	IV
HHI (seller) †	0.073	0.116		
	(0.171)	(0.172)		
HHI-WS (seller) †			0.199	-0.977**
			(0.219)	(0.445)
HHI-NS (seller) †			0.073	0.327*
			(0.171)	(0.190)
Capacity (log)	0.359***	0.351***	0.357***	0.337***
	(0.078)	(0.078)	(0.078)	(0.078)
Dummy (sister pulp mills)		0.423*		2.566***
		(0.237)		(0.836)
Dummy (UCFS capacity)	-0.533**	-0.555**	-0.541**	-0.582***
	(0.216)	(0.217)	(0.216)	(0.218)
Dummy (south)	2.636***	2.558***	2.585***	2.686***
	(0.313)	(0.316)	(0.318)	(0.324)
Constant	-0.912**	-1.021**	-0.933**	-1.364***
	(0.427)	(0.432)	(0.427)	(0.451)
Log Likelihood	-366.54	-364.94	-366.12	-361.22
N	648	648	648	648

Notes: †. Three-year lagged variable.

1. The dependent variable, vertical integration in capacity increase, is a dummy variable, indicating whether the increased capacity is vertically integrated or not compared to last year. The HHI of the seller market is used as a concentration measure for the chemical pulp market. HHI-WS denotes the HHI for the mill with sister mills in the neighbor market. HHI-NS denotes the HHI for an isolated mill in a neighbor market. These two variables are constructed by multiplying the HHI with relevant dummies. Dummy of sister pulp mills indicates whether a free sheet mill has a sister mill within its 350-mile radius circular neighbor market.

2. ***, **, *: the coefficient is significant at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.

Source: the FPL-UW data.

9. SUMMARY AND CONCLUSIONS

Based on the FPL-UW data, this study has produced results from cross-sectional, pooled, and panel logit models that analyze the relationship between market concentration and vertical integration in the pulp and paper industry. The transaction cost economics indicates that this relationship should be positive, which is verified in Ohanian (1994), as market concentration can be viewed as a proxy of asset specificity and transaction cost. However, the new updated data from FPL fail to repeat Ohanian's results in a duplicate model due to the existence of endogeneity, which is now revealed in that technological change tremendously changed the structure of industry.

Before extending our analysis to panel regression, due to the naturally integrated production nowadays of the low-grade paper product such as newsprint, kraft, and paperboard, we refine the model by narrowing the sample to free sheet mills and introducing a mill-specific measure for concentration. Free sheet mills are likely to purchase chemical pulp in the market because of higher profit margins even though chemical pulp is also much more expensive. With the introduction of a mill-specific concentration measure, we hope to locate the neighbor market in which the paper mill can purchase market pulp at a lower cost. Free sheet mills are further divided into two categories, those with or without sister pulp mills in their neighbor markets, because they have a different pattern of vertical integration at the mill level caused by the different status of vertical integration at the firm level.

In the cross-sectional and pooled logit models, the buyer concentration of paper mills is positively associated with their integration status, but the estimates are sometimes insignificant. Meanwhile, the coefficient of concentration for paper mills with access to in-firm purchasing is expected to be insignificant from zero; however, it is significantly negative in pooled models. Since controlling the technological factors of integration in cross-sectional models which benefits the production is difficult, we utilize the panel feature of the thirty-year data in order to control the time-invariant factors that affect the strategy of vertical integration and correlate with market concentration, such as forestland abundance. The estimates from the panel models all confirm our expectations. For isolated mills in neighbor pulp markets, market concentration is positively correlated with vertical integration. Furthermore, the mills along with sister pulp mills nearby are likely to get pulp within their own firm; thus, their strategy of integration is affected little by market concentration. Checked by various panel models, such as fixed effect, conditional fixed effect, and random effect models, the results are robust enough to support our hypothesis derived from transaction cost economics. Integration dynamics is investigated tentatively in a pooled logit model when we look at the strategy of vertical integration when free sheet mills increase capacity relative to the previous year. No dissimilarity between the models appears whether the focus is on vertical integration in total capacity or that in increased capacity.

In conclusion, many economic factors that affect vertical integration can be attributed to the technological economy, which should be controlled in empirical models. Previous studies did not thoroughly explore this particular industry, so they did not identify the real mechanism linking market concentration and vertical integration. We have redefined the measure for market concentration by narrowing our sample to mills that are most strongly affected by transaction cost in the paper industry. Then, using panel regression and robustness check, we successfully verify the sound relationship between transaction cost and vertical integration.

REFERENCES

Acemoglu, Daron, Philippe Aghion, Rachel Griffith, and Fabrizio Zilibotti, (2004), "Vertical Integration and Technology: Theory and Evidence," NBER Working Paper 10997, National Bureau of Economic Research.

Chamberlain, Gary, (1980), "Analysis of Covariance with Qualitative Data," *Review of Economic Studies*, 47(1): 225-38.

Diesen, Magnus, (1998), Economy of Integration, in Magnus Diesen (ed.), *Economics of the Pulp and Paper Industry*, Helsinki: Fapet Oy.

Globerman, Steven and Richard Schwindt, (1986), "The Organization of Vertically Related Transactions in the Canadian Forest Products Industries," *Journal of Economic Behavior and Organization*, 7(2): 199-212.

Greene, William, (2003), *Econometric Analysis*, 5th edition, New Jersey: Prentice-Hall.

Hsiao, Cheng, (2003), *Analysis of Panel Data*, 2nd edition, New York: Cambridge University Press.

Ince, Peter J., Xiaolei Li, Mo Zhou, Joseph Buongiorno, and Mary Reuter, (2001), "United States paper, paperboard, and market pulp capacity trends by process and location, 1970-2000," Research paper FPL-RP; 602, Forest Products Laboratory.

Joskow, Paul L., (2005), Vertical Integration, in Claude Ménard and Mary M. Shirley (eds.), *Handbook of New Institutional Economics*, Amsterdam: Springer.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian, (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, 21(2): 297-326.

Klein, Peter G., (2005), The Make-or-Buy Decision: Lessons from Empirical Studies, in Claude Ménard and Mary M. Shirley (eds.), *Handbook of New Institutional Economics*, Amsterdam: Springer.

MacDonald, James M., (1985), "Market Exchange or Vertical Integration: An Empirical Analysis," *Review of Economics and Statistics*, 67(2): 327-58.

Melendez, Marcela, (2002), *A Dynamic Model of Vertical Integration for the American Pulp and Paper Industry*, Ph.D. Dissertation, Yale University.

Ohanian, Nancy Kane, (1993), *The American pulp and paper industry, 1900-1940: Mill survival, firm structure, and industry relocation*, Westport, Conn. and London: Greenwood Press.

Ohanian, Nancy Kane, (1994), "Vertical Integration in the U.S. Pulp and Paper Industry, 1900-1940," *Review of Economics and Statistics*, 76(1): 202-07.

Toivanen, Hannes, (2004), *Learning and Corporate Strategy: The Dynamic Evolution of the North American Pulp and Paper Industry, 1860-1960*, Ph.D. Dissertation, Georgia Institute of Technology.

Whinston, Michael D., (2003), "On the Transaction Cost Determinants of Vertical Integration," *Journal of Law Economics, and Organization*, 19(1): 1-23.

Williamson, Oliver E., (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: The Free Press.

Zavatta, Roberto, (1993), The Pulp and Paper Industry, in H. W. de Jong (ed.), *The Structure of European Industry*, 3rd edition, Amsterdam: Kluwer Academic Publishers.